Data Dictionary – SQL view–derived fields (clean/BI layer)

This document lists engineered fields created in SQL views for analysis and Tableau consumption.

Field Name	Description
pickup_hour	Pickup hour of day (0–23) derived from tpep_pickup_datetime.
pickup_dow	Pickup day-of-week as integer derived from tpep_pickup_datetime.
pickup_date	Pickup date (date part of tpep_pickup_datetime).
pickup_month	Pickup month derived from tpep_pickup_datetime (for month-level aggregation).
duration_min	Trip duration in minutes (tpep_dropoff_datetime – tpep_pickup_datetime).
pretip_total	Total amount before tip: total_amount – tip_amount.
extras_pretip	Non-fare add-ons before tip: pretip_total – fare_amount (fees/extras proxy).
tip_rate_fare	Tip rate using fare as denominator: tip_amount / fare_amount.
tip_rate_pretip	Tip rate using pre-tip total as denominator: tip_amount / pretip_total.
is_tipped	Binary tipping flag: 1 if tip_amount > 0 else 0.
tip_rate_pretip_capped	tip_rate_pretip capped at a high percentile to reduce outlier distortion.
tip_rate_fare_capped	tip_rate_fare capped at a high percentile to reduce outlier distortion.
payment_group	Payment type bucket derived from payment_type (e.g., Card/Cash/Other).
duration_per_mile	Minutes per mile: duration_min / trip_distance (congestion/stress proxy).
is_airport_trip	Airport context flag used for airport vs non-airport segmentation.
pu_zone	Pickup zone name joined from the taxi zone dimension (PULocationID).
do_zone	Dropoff zone name joined from the taxi zone dimension (DOLocationID).
pu_borough	Pickup borough joined from the taxi zone dimension (PULocationID).
do_borough	Dropoff borough joined from the taxi zone dimension (DOLocationID).
pu_has_bid	0/1 flag: pickup zone overlaps at least one BID polygon.
do_has_bid	0/1 flag: dropoff zone overlaps at least one BID polygon.
pu_bid_count	Number of distinct BIDs overlapping the pickup taxi zone.
do_bid_count	Number of distinct BIDs overlapping the dropoff taxi zone.
pu_bid_overlap_share	Share of pickup taxi zone area covered by BID polygons (0–1).
do_bid_overlap_share	Share of dropoff taxi zone area covered by BID polygons (0–1).
pu_bid_overlap_share_pct	pu_bid_overlap_share expressed as percent (0–100).
do_bid_overlap_share_pct	do_bid_overlap_share expressed as percent (0–100).